Diffused Heads: Diffusion Models Beat GANs on Talking-Face Generation

Year: 2023 | Citations: 174 | Authors: Michal Stypulkowski, Konstantinos Vougioukas, Sen He, Maciej Zi■ba, Stavros Petridis

Abstract

Talking face generation has historically struggled to produce head movements and natural facial expressions without guidance from additional reference videos. Recent developments in diffusion-based generative models allow for more realistic and stable data synthesis and their performance on image and video generation has surpassed that of other generative models. In this work, we present an autoregressive diffusion model that requires only one identity image and audio sequence to generate a video of a realistic talking head. Our solution is capable of hallucinating head movements, facial expressions, such as blinks, and preserving a given background. We evaluate our model on two different datasets, achieving state-of-the-art results in expressiveness and smoothness on both of them.1